

Roark

#2

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/726,899

DATE: 01/27/2001
TIME: 00:25:33

INPUT SET: S36328.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information
4
5 (i) APPLICANT: Bandman, Olga
6 Goli, Surya K.
7 Hillman, Jennifer L.
8
9 (ii) TITLE OF THE INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
10
11 (iii) NUMBER OF SEQUENCES: 12
12
13 (iv) CORRESPONDENCE ADDRESS:
14 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
15 (B) STREET: 3174 Porter Drive
16 (C) CITY: Palo Alto
17 (D) STATE: CA
18 (E) COUNTRY: USA
19 (F) ZIP: 94304
20
21 (v) COMPUTER READABLE FORM:
22 (A) MEDIUM TYPE: Diskette
23 (B) COMPUTER: IBM Compatible
24 (C) OPERATING SYSTEM: DOS
25 (D) SOFTWARE: FastSEQ for Windows Version 2.0
26
27 (vi) CURRENT APPLICATION DATA:
28 (A) APPLICATION NUMBER: 09/726,899
29 (B) FILING DATE:
30 (C) CLASSIFICATION:
31
32 (vii) PRIOR APPLICATION DATA:
33 (A) APPLICATION NUMBER: 08/785,065
34 (B) FILING DATE:
35
36 (viii) ATTORNEY/AGENT INFORMATION:
37 (A) NAME: Billings, Lucy J.
38 (B) REGISTRATION NUMBER: 36,749
39 (C) REFERENCE/DOCKET NUMBER: PF-0187 US
40
41 (ix) TELECOMMUNICATION INFORMATION:
42 (A) TELEPHONE: 415-855-0555
43 (B) TELEFAX: 415-845-4166
44 (C) TELEX:
45
46

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/726,899DATE: 01/27/2001
TIME: 00:25:33

INPUT SET: S36328.raw

47 (2) INFORMATION FOR SEQ ID NO:1:
48

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 264 amino acids

51 (B) TYPE: amino acid

52 (C) STRANDEDNESS: single

53 (D) TOPOLOGY: linear
54

55 (vii) IMMEDIATE SOURCE:

56 (A) LIBRARY: Consensus

57 (B) CLONE: Consensus
5859 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
60

61	Met	Ala	Ala	Ala	Ala	Val	Ala	Arg	Leu	Trp	Trp	Arg	Gly	Ile	Leu	Gly
62	1				5					10					15	
63	Ala	Ser	Ala	Leu	Thr	Arg	Gly	Thr	Gly	Arg	Pro	Ser	Val	Leu	Leu	Leu
64				20					25					30		
65	Pro	Val	Arg	Arg	Glu	Ser	Ala	Gly	Ala	Asp	Thr	Arg	Pro	Thr	Val	Arg
66			35					40					45			
67	Pro	Arg	Asn	Asp	Val	Ala	His	Lys	Gln	Leu	Ser	Ala	Phe	Gly	Glu	Tyr
68		50					55					60				
69	Val	Ala	Glu	Ile	Leu	Pro	Lys	Tyr	Val	Gln	Gln	Val	Gln	Val	Ser	Cys
70	65				70				75						80	
71	Phe	Asn	Glu	Leu	Glu	Val	Cys	Ile	His	Pro	Asp	Gly	Val	Ile	Pro	Val
72				85					90					95		
73	Leu	Thr	Phe	Leu	Arg	Asp	His	Thr	Asn	Ala	Gln	Phe	Lys	Ser	Leu	Val
74				100					105					110		
75	Asp	Leu	Thr	Ala	Val	Asp	Val	Pro	Thr	Arg	Gln	Asn	Arg	Phe	Glu	Ile
76			115					120					125			
77	Val	Tyr	Asn	Leu	Leu	Ser	Leu	Arg	Phe	Asn	Ser	Arg	Ile	Arg	Val	Lys
78		130					135					140				
79	Thr	Tyr	Thr	Asp	Glu	Leu	Thr	Pro	Ile	Glu	Ser	Ala	Val	Ser	Val	Phe
80	145				150					155					160	
81	Lys	Ala	Ala	Asn	Trp	Tyr	Glu	Arg	Glu	Ile	Trp	Asp	Met	Phe	Gly	Val
82				165					170					175		
83	Phe	Phe	Ala	Asn	His	Pro	Asp	Leu	Arg	Arg	Ile	Leu	Thr	Asp	Tyr	Gly
84				180					185					190		
85	Phe	Glu	Gly	His	Pro	Phe	Arg	Lys	Asp	Phe	Pro	Leu	Ser	Gly	Tyr	Val
86			195					200					205			
87	Glu	Leu	Arg	Tyr	Asp	Asp	Glu	Val	Lys	Arg	Val	Val	Ala	Glu	Pro	Val
88		210					215					220				
89	Glu	Leu	Ala	Gln	Glu	Phe	Arg	Lys	Phe	Asp	Leu	Asn	Ser	Pro	Trp	Glu
90	225				230					235					240	
91	Ala	Phe	Pro	Val	Tyr	Arg	Gln	Pro	Pro	Glu	Ser	Leu	Lys	Leu	Glu	Ala
92				245						250					255	
93	Gly	Asp	Lys	Lys	Pro	Asp	Ala	Lys								
94				260												

95
96 (2) INFORMATION FOR SEQ ID NO:2:
97

98 (i) SEQUENCE CHARACTERISTICS:

99 (A) LENGTH: 1023 base pairs

RAW SEQUENCE LISTING PATENT APPLICATION US/09/726,899

DATE: 01/27/2001
TIME: 00:25:33

INPUT SET: S36328.raw

100 (B) TYPE: nucleic acid
101 (C) STRANDEDNESS: single
102 (D) TOPOLOGY: linear

103
104 (vii) IMMEDIATE SOURCE:
105 (A) LIBRARY: Consensus
106 (B) CLONE: Consensus
107

108 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
109

110	GA	AACTCTAAT	ACGAGCACTA	TAGGGAAAGC	TGGTAGCCTG	CAGGTACCGG	TCCGGAATTC	60
111	CC	GGGTCGAC	CCACGCGTCC	GCCGTGCCCT	TGGGGCTCCG	TGTCCTGCTG	TCTTTCCGTC	120
112	CG	CTGCCTAG	TCTGCATCTG	AGTAACATGG	CGGCGGCGGC	GGTAGCCAGG	CTGTGGTGGC	180
113	GCG	GATCTT	GGGGGCCTCG	GCGCTGACCA	GGGGGACTGG	GCGACCCCTCC	GTTCTGTTGC	240
114	TGCC	GGTGAG	GCGGGAGAGC	GCCGGGGCCG	ACACGCGCCC	CACTGTCAGA	CCACGGAATG	300
115	ATGT	GGCCCA	CAAGCAGCTC	TCAGCTTTTG	GAGAGTATGT	GGCTGAAATC	TTGCCCAAGT	360
116	ATGT	CCAACA	AGTTCAGGTG	TCCTGCTTCA	ATGAGTTAGA	GGTCTGTATC	CATCCTGATG	420
117	GCGT	CATCCC	AGTGCTGACT	TTCTCTCAGG	ATCACACCAA	TGCACAGTTC	AAATCTCTGG	480
118	TTGA	CTTGAC	AGCAGTGGAC	GTCCCAACTC	GGCAAACCCG	TTTTGAGATT	GTCTACAACC	540
119	TGTT	TGTCTCT	GCGCTTCAAC	TCACGGATCC	GTGTGAAGAC	CTACACAGAT	GAGCTGACGC	600
120	CCATT	GAGTC	TGCTGTCTCT	GTGTTCAAGG	CAGCCAACTG	GTATGAAAGG	GAGATCTGGG	660
121	ACAT	GTTTGG	AGTCTTCTTT	GCTAACCACC	CTGATCTAAG	AAGGATCCTG	ACAGATTATG	720
122	GCTTC	GAGGG	ACATCCTTTC	CGGAAAGACT	TTCTCTATC	TGGCTATGTT	GAGTTACGTT	780
123	ATGA	TGATGA	AGTGAAGCGT	GTGGTGGCAG	AGCCGGTGGA	GTTGGCCCAA	GAGTTCCGCA	840
124	AATTT	GACCT	GAACAGCCCC	TGGGAGGCTT	TCCCAGTCTA	TCGCCAACCC	CCGGAGAGTC	900
125	TCAAG	CTTGA	AGCCGGAGAC	AAGAAGCCTG	ATGCCAAGTA	GCTCCAGGGA	ACGCATGTGG	960
126	ATCCT	AGACA	GCGCCTTATC	TATGATTGAG	TGTCCGTGTA	AATAAATTCC	TACTTAGACT	1020
127	TAC							1023

128
129 (2) INFORMATION FOR SEQ ID NO:3:
130

131 (i) SEQUENCE CHARACTERISTICS:
132 (A) LENGTH: 129 amino acids
133 (B) TYPE: amino acid
134 (C) STRANDEDNESS: single
135 (D) TOPOLOGY: linear
136

137 (vii) IMMEDIATE SOURCE:
138 (A) LIBRARY: Consensus
139 (B) CLONE: Consensus
140

141 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
142

143	Met	Ser	Phe	Pro	Lys	Tyr	Lys	Pro	Ser	Ser	Leu	Arg	Thr	Leu	Pro	Glu
144	1				5					10					15	
145	Thr	Leu	Asp	Pro	Ala	Glu	Tyr	Asn	Ile	Ser	Pro	Glu	Thr	Arg	Arg	Ala
146				20					25					30		
147	Gln	Ala	Glu	Arg	Leu	Ala	Ile	Arg	Ala	Gln	Leu	Lys	Arg	Glu	Tyr	Leu
148			35					40					45			
149	Leu	Gln	Tyr	Asn	Asp	Pro	Asn	Arg	Arg	Gly	Leu	Ile	Glu	Asn	Pro	Ala
150		50					55					60				
151	Leu	Leu	Arg	Trp	Ala	Tyr	Ala	Arg	Thr	Ile	Asn	Val	Tyr	Pro	Asn	Phe
152	65					70				75					80	

RAW SEQUENCE LISTING PATENT APPLICATION US/09/726,899

DATE: 01/27/2001
TIME: 00:25:34

INPUT SET: S36328.raw

153 Arg Pro Thr Pro Lys Asn Ser Leu Met Gly Ala Leu Cys Gly Phe Gly
154 85 90 95
155 Pro Leu Ile Phe Ile Tyr Tyr Ile Ile Lys Thr Glu Arg Asp Arg Lys
156 100 105 110
157 Glu Lys Leu Ile Gln Glu Gly Lys Leu Asp Arg Thr Phe His Leu Ser
158 115 120 125
159 Tyr
160
161

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 451 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

176	CCAAGATGTC	GTTCCCAAAG	TATAAGCCGT	CGAGCCTGCG	CACTCTGCCT	GAGACCCTCG	60
177	ACCCAGCCGA	ATACAACATA	TCTCCGGAAA	CCCGGCGGGC	GCAAGCGGAG	CGGTTGGCCA	120
178	TAAGAGCCCA	GCTGAAACGA	GAGTACCTGC	TTCAGTACAA	CGATCCCAAC	CGCCGAGGGC	180
179	TCATCGAAAA	TCCTGCCTTG	CTTCGTTGGG	CCTATGCAAG	AACAATAAAT	GTCTATCCTA	240
180	ATTTTCAGACC	CACTCCTAAA	AACTCACTCA	TGGGAGCTCT	GTGTGGATTT	GGGCCCCTCA	300
181	TCTTCATTTA	TTATATTATC	AAAAC TGAGA	GGGATAGGAA	AGAAAAACTT	ATCCAGGAAG	360
182	GAAAATTGGA	TCGAACATTT	CACCTCTCAT	ATTAAGTCTG	GCAATGATGA	CTATATGTAT	420
183	TCCTGCCTAA	ATAAATCATC	TATTAATCAT	T			451

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 106 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

199	Met	Pro	Phe	Leu	Asp	Ile	Gln	Lys	Arg	Phe	Gly	Leu	Asn	Ile	Asp	Arg
200	1				5				10						15	
201	Trp	Leu	Thr	Ile	Gln	Ser	Gly	Glu	Gln	Pro	Tyr	Lys	Met	Ala	Gly	Arg
202				20				25						30		
203	Cys	His	Ala	Phe	Glu	Lys	Glu	Trp	Ile	Glu	Cys	Ala	His	Gly	Ile	Gly
204			35				40				45					
205	Tyr	Thr	Arg	Ala	Glu	Lys	Glu	Cys	Lys	Ile	Glu	Tyr	Asp	Asp	Phe	Val

RAW SEQUENCE LISTING PATENT APPLICATION US/09/726,899

DATE: 01/27/2001
TIME: 00:25:34

INPUT SET: S36328.raw

```

206          50          55          60
207 Glu Cys Leu Leu Arg Gln Lys Thr Met Arg Arg Ala Gly Thr Ile Arg
208 65          70          75          80
209 Lys Gln Arg Asp Lys Leu Ile Lys Glu Gly Lys Tyr Thr Pro Pro Pro
210          85          90          95
211 His His Ile Gly Lys Gly Glu Pro Arg Pro
212          100          105
213

```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 470 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

227
228 AGCTAGTCGT TCTGAAGCGG CGGCCAGAGA AGAGTCAAGG GCACGAGCAT CGGCCATGCC      60
229 TTTCTTGGAC ATCCAGAAAA GGTTCCGGCCT TAACATAGAT CGATGGTTGA CAATCCAGAG      120
230 TGGTGAACAG CCCTACAAGA TGGCTGGTTCG ATGCCATGCT TTTGAAAAAG AATGGATAGA      180
231 ATGTGCACAT GGAATCGGTT ATACTCGGGC AGAGAAAGAG TGCAAGATAG AATATGATGA      240
232 TTTCTGTAGAG TGTTTGCTTC GGCAGAAAAC GATGAGACGT GCAGGTACCA TCAGGAAGCA      300
233 GCGGGATAAG CTGATAAAGG AAGGAAAGTA CACCCCTCCA CCTCACCACA TTGGCAAGGG      360
234 GGAGCCTCGG CCCTGAACAG AGCAGCTGCT GATGTCTGGA GGCTGATTTT CCTGTTCTCT      420
235 GTTCTCCACT GGAAAGGTTG TTTACGACAA ACCTCCTTGT CAAAGTGTGT      470
236

```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

250
251 Met Ile Ala Arg Arg Asn Pro Glu Pro Leu Arg Phe Leu Pro Asp Glu
252 1          5          10          15
253 Ala Arg Ser Leu Pro Pro Pro Lys Leu Thr Asp Pro Arg Leu Leu Tyr
254          20          25          30
255 Ile Gly Phe Leu Gly Tyr Cys Ser Gly Leu Ile Asp Asn Leu Ile Arg
256          35          40          45
257 Arg Arg Pro Ile Ala Thr Ala Gly Leu His Arg Gln Xaa Xaa Tyr Ile
258 50          55          60

```

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/726,899

DATE: 01/27/2001
TIME: 00:25:35

INPUT SET: S36328.raw

Line	Error	Original Text
------	-------	---------------

PAGE: 1

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION US/09/726,899

DATE: 01/27/2001
TIME: 00:25:35

INPUT SET: S36328.raw

< < THERE ARE NO ITEMS MISSING > >

PAGE: 1

SEQUENCE CORRECTION REPORT
PATENT APPLICATION US/09/726,899

DATE: 01/27/2001
TIME: 00:25:35

INPUT SET: S36328.raw

Line	Original Text	Corrected Text
3	(1) General Information	(1) GENERAL INFORMATION:
9	(ii) TITLE OF THE INVENTION: NOVEL SUBUNITS O	(ii) TITLE OF INVENTION: NOVEL SUBUNITS OF NA